Improving Voice and Data Services in Cellular/WLAN Integrated Networks by Admission Control

Wei Song, *Student Member, IEEE*, Yu Cheng, *Member, IEEE*, and Weihua Zhuang, *Senior Member, IEEE*

***Abstract*— In this paper, we study voice and data service provisioning in an integrated system of cellular and wireless local area networks (WLANs). With the ubiquitous coverage of the cellular network and the disjoint deployment of WLANs in hot-spot areas, the integrated system has a two-tier overlaying structure. As an essential resource allocation aspect, admission control can be used to properly admit voice and data calls to the overlaying cells and WLANs. A simple admission scheme is proposed in this study to analyze the dependence of resource utilization and the impact of user mobility and traffic characteristics on admission parameters. Both admission control and rate control are considered to limit the input traffic to the WLAN, so that the WLAN operates in its most efficient states and effectively complements the cellular network. The call blocking/dropping probabilities and data call throughput are evaluated for effective and accurate derivation of the admission parameters. It is observed that the utilization varies with the configuration of admission parameters, which properly distributes the voice and data traffic load to the cells and WLANs. Mobility and traffic variability have a significant impact on the selection of the admission parameters.**

***Index Terms*— Cellular/WLAN interworking, quality of service, call admission control, rate control, WLAN capacity.**

I. INTRODUCTION

WITH complementary characteristics presented by the cellular network and wireless local area network (WLAN), their interworking becomes a promising trend for next-generation wireless networks. In a cellular/WLAN integrated network, the cellular network provides ubiquitous coverage, while the WLANs are deployed disjointly in hot-spot areas. Then, in the area with WLAN coverage (referred to as *double-coverage area* in the following), both cellular and WLAN accesses are available. With this two-tier overlaying structure, a new call in the double-coverage area can be admitted either to a cell or to a WLAN. Moreover, proper admission decision is needed for vertical handoffs between a cell and a WLAN, which may not be necessary to maintain a call, but is mainly for load balancing and/or quality of service (QoS) enhancement. The heterogeneous underlying QoS support of the integrated network can be exploited to maximize the overall resource utilization. In the cellular network, the base station (BS) controls access to the shared radio spectrum and reserves resources for admitted calls. This centralized control and reservation-based resource allocation enables fine-grained QoS. On the other hand, the medium access control (MAC) in WLANs is usually a contention-based random access protocol, e.g., the distributed coordination function (DCF) of IEEE 802.11. Because of inevitable collision and backoff, this type of MAC is difficult to support services with strict QoS requirements such as real-time voice service, although it is efficient in serving bursty data traffic. With this complementary QoS support, when multiple services are considered, the traffic load of different classes should be properly distributed to the cells and WLANs by admission control.

In [1], we analyze the performance of a simple admission scheme for voice and data services in the cellular/WLAN integrated network, referred to as the *WLAN-first scheme*. In the scheme, both new voice and data calls first request admission to the WLAN and, if rejected, overflow to the cell. Ongoing voice and data calls try to hand over to the WLAN whenever users move into the WLAN coverage. In contrast, the service-differentiated admission scheme proposed in [2] applies a different admission strategy for voice service, in which the cellular network is the first choice for voice calls and no vertical handoff from the cell to the WLAN is executed for ongoing voice calls. To maximize resource utilization, a complex set of admission parameters in both schemes needs to be determined so that the voice and data traffic load is properly distributed to the cells and WLANs. In this research, we further simplify the admission strategies and generalize the above admission schemes. As such, the dependence of resource utilization on admission parameters is investigated more thoroughly. Secondly, we extend the QoS evaluation approaches in our previous works [3, 4] so as to more effectively derive the admission parameters. In [3], the QoS metrics are evaluated by two-dimensional Markov chains, which is very accurate and can well take into account the high variability of data traffic [5] and user mobility within the WLAN [6]. Unfortunately, the computation complexity increases with the size of the state space. In [4], the QoS evaluation is based on moment generating functions (MGFs), which is more efficient but at the price of accuracy. Moreover, the user mobility within the WLAN is not differentiated from that of other areas, and the high variability of data traffic is

Manuscript received April 6, 2006; revised September 7, 2006 and December 7, 2006; accepted January 23, 2007. The associate editor coordinating the review of this paper and approving it for publication was V. Leung.

W. Zhuang and W. Song are with the Centre for Wireless Communications, Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1 (email: {wzhuang, wsong}@bbcr.uwaterloo.ca).

Y. Cheng is with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, 3301 South Dearborn Street, Chicago, IL, USA 60616-3793 (email: cheng@iit.edu).

Digital Object Identifier 10.1109/TWC.2007.060146.

neglected for simplicity. In this paper, we extend the MGF-based approach to take into account the high variability of user mobility and data traffic.

It is observed in [1] that the system performance closely depends on the characteristics of the WLAN in supporting different services. A high utilization is achieved when the resource allocation by admission control matches well the system characteristics and keeps the WLAN operating in its most efficient states. In [1], the data traffic is assumed to be saturated because, for data applications such as Web browsing, the data file to be transmitted is usually pre-stored in a server. However, it is found in [7] that, when the number of contention nodes is large enough, the WLAN channel utilization is maximized in the unsaturated case instead of the saturated case. Hence, in this study, we consider applying rate control to regulate downlink data traffic at the access point [8]. The traffic regulation can take good advantage of the elasticity of data traffic to improve the voice support of the WLAN; at the same time, it allows more data calls admitted to the WLAN and more effectively complements the data service provisioning of the cellular network.

The remainder of this paper is organized as follows. We describe in Section II the user mobility, voice/data traffic, the WLAN and cellular capacity. Section III discusses the admission control problem for cellular/WLAN interworking and the QoS evaluation approach to determine admission parameters. Numerical results are presented in Section IV, and Section V concludes this research.

II. System Model

We consider an interworking scenario with a simple topology, in which there is one WLAN located in each cell, referred to as a *cell/WLAN cluster*. Within the interworked system, location-dependent factors such as user mobility, traffic distribution, and underlying network support need to be taken into account. The important symbols used in this paper are summarized in Table I for reference.

A. Location-Dependent Mobility Model

As shown in [6], the indoor deployment and low user mobility within WLANs results in heavy-tailed user residence time within a WLAN, denoted by T_r^w. As performance analysis becomes extremely difficult with heavy-tailed distributions, certain approximation is necessary to make the analysis effective and tractable. It is proposed in [9] to approximate a large class of heavy-tailed distributions (including Pareto and Weibull distributions) with hyper-exponential distributions. An important feature of heavy-tailedness is the so-called "*mice-elephants*" phenomenon [10]. For the WLAN residence time, it implies that most users stay within a WLAN for a short time, while a small fraction of the users have an extremely long residence time. To explore its impact on performance, we adopt a two-stage hyper-exponential distribution, which is very simple and also well captures the high variability. The probability density function (PDF) of T_r^w with mean $1/\eta^w$ is then given by

$$f_{T_r^w}(t) = \frac{a}{a+1} \cdot \frac{1}{\frac{1}{a} \cdot \frac{1}{\eta^w}} e^{-a\eta^w t} + \frac{1}{a+1} \cdot \frac{1}{a \cdot \frac{1}{\eta^w}} e^{-\frac{\eta^w}{a} t}$$
$$a \geq 1, \qquad t \geq 0. \tag{1}$$

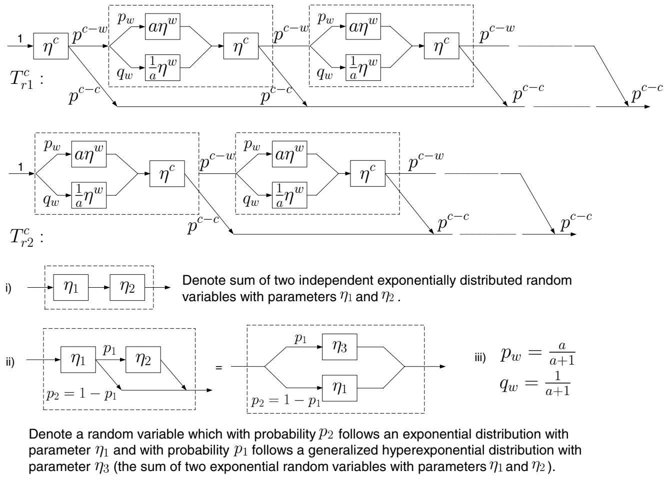

Fig. 1. Modeling of user residence time in a cell.

A large fraction $\frac{a}{a+1}$ of the users stay within the WLAN for a mean time $\frac{1}{a} \cdot \frac{1}{\eta^w}$, while the other $\frac{1}{a+1}$ of the users have a mean residence time of $a \cdot \frac{1}{\eta^w}$. The coefficient of variance of T_r^w is $C_{v,T_r^w} = \sqrt{2a + (2/a) - 3}$. Increasing the parameter a results in T_r^w with a higher variability. Since the hyper-exponential distribution consists of a linear mixture of exponentials, the following analysis can be extended to consider hyper-exponential distributions with more exponential components, which can more accurately approximate the original heavy-tailed distribution.

On the other hand, the user residence time in the area of a cell with only cellular access (referred to as *cellular-only area*), denoted by T_r^c, is assumed to be exponentially distributed with parameter η^c. Users moving out of the cellular-only area enter neighboring cells with a probability p^{c-c} and enter the coverage of the overlaying WLAN in the target cell with a probability $p^{c-w} = 1 - p^{c-c}$. As a result, the residence time of users admitted in the cell follows a more complicate phase-type distribution as shown in Fig. 1. Let T_{r1}^c and T_{r2}^c denote the residence time within the cell of a new or handoff call admitted to the cell from the cellular-only area and that of a new call admitted to the cell from the double-coverage area, respectively. The MGFs of T_{r1}^c and T_{r2}^c are derived as

$$\Phi_1(s) = \sum_{i=1}^{\infty} (p^{c-w})^{i-1} p^{c-c} \frac{\eta^c}{\eta^c - s} [\psi(s)]^{i-1} \tag{2}$$

$$\Phi_2(s) = \sum_{i=1}^{\infty} (p^{c-w})^{i-1} p^{c-c} [\psi(s)]^{i} \tag{3}$$

where $\psi(\cdot)$ is the MGF of $(T_r^c + T_r^w)$, given by

$$\psi(s) = \mathrm{E}\left[e^{s(T_r^c + T_r^w)}\right] \tag{4}$$
$$= \frac{\eta^c}{\eta^c - s} \left[\frac{a}{a+1} \cdot \frac{a\eta^w}{a\eta^w - s} + \frac{1}{a+1} \cdot \frac{\frac{1}{a}\eta^w}{\frac{1}{a}\eta^w - s} \right].$$

B. Traffic Models for Voice and Data Services

As an important motivation for cellular/WLAN interworking, multi-service support is an essential requirement for future wireless networks. For example, four service classes are defined in the specification 3GPP TS 23.107 for the universal

TABLE I
SUMMARY OF IMPORTANT SYMBOLS

Symbol	Definition
a (b)	User mobility (data traffic) variability parameter
B_v^c (B_d^c)	Blocking/dropping probability of a cell for voice (data) calls
B_v^w (B_d^w)	Blocking probability of a WLAN for voice (data) calls
W_c	Cell bandwidth
D_v^w (D_d^w)	Rejection probability of a WLAN for handoff voice (data) calls from a cell
f_d	Mean data call size
G_v^w (G_d^w)	Number of guard channels for new voice (data) calls to a WLAN
N_v^c (N_d^c)	Maximum number of voice (data) calls allowed in a cell
N_v^w (N_d^w)	Maximum number of voice (data) calls allowed in a WLAN
p^{c-c} (p^{c-w})	Probability of a user moving out of current cell to neighboring cells (to overlaying WLAN)
$S_v(\cdot)$ ($S_d(\cdot)$)	Service time of packets from a voice (data) flow in a WLAN
T_{r1}^c (T_{r2}^c)	Residence time within a cell of a call admitted to a cell from cellular-only (double-coverage) area
T_r^c (T_r^w)	User residence time in cellular-only (double-coverage) area of a cell with mean $(\eta^c)^{-1}((\eta^w)^{-1})$
T_v	Voice call duration with mean $(\mu_v)^{-1}$
$T_v^p(\cdot)$	Voice packet delay in a WLAN
λ_{d1} (λ_{d2})	Mean new data call arrival rate in cellular-only (double-coverage) area of a cell
λ_{hv}^{c-c} (λ_{hd}^{c-c})	Mean arrival rate of handoff voice (data) calls between neighboring cells
λ_{hv}^{c-w} (λ_{hd}^{c-w})	Mean arrival rate of handoff voice (data) calls from cell to overlaying WLAN
λ_{hv}^{w-c} (λ_{hd}^{w-c})	Mean arrival rate of handoff voice (data) calls from WLAN to overlaying cell
λ_{nv2}^c (λ_{nd2}^c)	Mean arrival rate of new voice (data) calls to a cell from double-coverage area
λ_{nv}^w (λ_{nd}^w)	Mean arrival rate of new voice (data) calls to a WLAN
λ_{v1} (λ_{v2})	Mean new voice call arrival rate in cellular-only (double-coverage) area of a cell
λ_v^p ($\lambda_d^p(\cdot)$)	Mean packet arrival rate from a voice (data) flow
$\pi_v^c(\cdot)$ ($\pi_d^c(\cdot)$)	Steady-state probability of the number of voice (data) calls in a cell
$\pi_v^w(\cdot)$ ($\pi_d^w(\cdot)$)	Steady-state probability of the number of voice (data) calls in a WLAN
θ_v^c (θ_d^c)	Probability that a voice (data) call in the double-coverage area requests admission to a cell
θ_v^w (θ_d^w)	Probability that a voice (data) call in the double-coverage area requests admission to a WLAN
$\xi_v^w(\cdot)$ ($\xi_d^w(\cdot)$)	Mean packet service rate for a voice (data) flow (frames/slot)
$\psi(\cdot)$, $\Phi_1(\cdot)$, $\Phi_2(\cdot)$	Moment generating functions of $T_r^c + T_r^w, T_{r1}^c$, and T_{r2}^c, respectively
Γ_d^c (Γ_d^w)	Mean data call throughput in a cell (WLAN)

mobile telecommunication system (UMTS). In this study, we consider both voice (such as traditional telephony) and data (such as Web browsing) services, which are a representative of the conversational and interactive service classes of UMTS, respectively. The call-level QoS requirements are defined in terms of voice call blocking/dropping probability (Q_{vb}), data call blocking/dropping probability (Q_{db}), and data call throughput (Q_{dt}). As widely assumed in previous works, voice and data call arrivals are independent Poisson processes. The voice call duration T_v is exponentially distributed with mean $1/\mu_v$. It is observed in [5] that the data traffic generated by Web browsing is bursty and one main reason for the high variability is the heavy-tailed file size. To explore the impact of highly variable file size on performance, similar to the modeling of WLAN residence time, the data call size (L_d) is modeled by a two-stage hyper-exponential distribution with mean f_d and PDF given by

$$f_{L_d}(x) = \frac{b}{b+1} \cdot \frac{1}{\frac{1}{b} \cdot f_d} e^{-\frac{b}{f_d} x} + \frac{1}{b+1} \cdot \frac{1}{b \cdot f_d} e^{-\frac{1}{b} \cdot \frac{1}{f_d} x}$$
$$b \geq 1, \qquad x \geq 0. \tag{5}$$

A larger value of b corresponds to a data call size with a higher variability.

In the cellular network, with the reservation-based resource allocation, the restricted access mechanism [11] can be applied for resource sharing between voice and data in the cell. A certain bandwidth can be reserved for each voice call to guarantee its packet-level requirements such as packet delay. Given the total cell bandwidth W_c, N_v^c voice calls at maximum are accommodated to meet the requirement on voice call blocking/dropping probability, while the other bandwidth is dedicated to data. Moreover, the elasticity of data traffic is exploited by allowing ongoing data calls to equally share all bandwidth unused by current voice traffic. Considering the interactive nature of some data applications, each data call needs to finish transfer within a time limit and be guaranteed certain throughput. As a result, the number of data calls admitted in the cell is limited by N_d^c.

C. WLAN Model for Capacity Analysis

With the contention-based access, the WLAN resources are shared in a complete sharing manner and there is no QoS guarantee for the delay or throughput requirements of voice and data calls. It is necessary to apply admission control to restrict the admitted calls contending for access. In this study, we consider both two-way voice calls and one-way downlink data calls such as the download of a Web page or transfer of a data file. Let N_v^w and N_d^w denote the maximum numbers of voice and data calls allowed in the WLAN, respectively. Each voice call has two flows from and to the mobile station (MS), while each data call has one flow to the MS. Since different QoS support is provided by the WLAN to voice and data calls having different traffic characteristics and QoS requirements, the WLAN utilization changes with the amounts of voice and data traffic sharing the WLAN channel. To ensure that the WLAN operates efficiently and complements the cellular network effectively, N_v^w and N_d^w should be properly selected within the WLAN capacity region, which is defined in terms of the maximum numbers (n_v^w, n_d^w) of voice and data flows simultaneously served in the WLAN.

It is observed in [7] that there exists an optimal operating point for the WLAN in the unsaturated case, beyond which the packet delay increases dramatically and the throughput drops quickly. When the packet service rate is larger than the arrival rate (unsaturated case) and the collision probability is small enough (e.g., less than 0.1), the service queue of a flow is almost empty and the packet delay (say, less than 30 ms) is sufficiently small to meet the requirement of the real-time voice service [7]. Hence, we consider unsaturated downlink data traffic, which is regulated by applying rate control at the access point [8]. The voice traffic is also considered unsaturated depending on the codec. Let λ_v^p denote the mean rate (in frames/slot) of packet arrivals from a voice source. For simplicity, the voice and data packet sizes are assumed to be constant and the packet arrivals from a voice or data flow are assumed to be a Poisson process. To admit n_v^w voice flows and n_d^w data flows in the WLAN, the data packet input rate $\lambda_d^p(n_v^w, n_d^w)$ should be limited so as to keep a small collision probability and satisfy the network stability constraints, i.e., $\lambda_v^p < \xi_v^w(n_v^w, n_d^w)$ and $\lambda_d^p(n_v^w, n_d^w) < \xi_d^w(n_v^w, n_d^w)$, where $\xi_v^w(n_v^w, n_d^w)$ and $\xi_d^w(n_v^w, n_d^w)$ are the average service rates for packets from one voice and data flow, respectively. The maximum $\lambda_d^p(n_v^w, n_d^w)$ and the corresponding $\xi_v^w(n_v^w, n_d^w)$ and $\xi_d^w(n_v^w, n_d^w)$ that satisfy the above constraints can be determined recursively using the following approach.

To analyze the servicing process for different traffic flows, we can decompose the WLAN modeling into two parts. First, the service at each flow is modeled by an $M/G/1$ queueing system. Each packet from a flow to be transmitted is buffered in the queue and then contends for channel access. Only when the packet departs can the next packet moves to the queue head. Second, from the perspective of the WLAN channel, packets from different flows are served in a contention-based access manner, which determines the packet service time, i.e., the time from the instant that a packet becomes the head of the flow's service queue to the time that it is successfully transmitted or discarded after reaching retransmission limit. Given the moments of the packet service time, the average packet delay can be obtained from the Pollaczek-Khinchine (P-K) formula [12]. In [13], the packet transmission process is modeled with a Markov chain. By using the signal transfer function of the corresponding state transition diagram, the probability generating function (PGF) of the packet service time is derived. Then, the arbitrary n^{th} moment of the packet service time can be obtained by differentiation. However, the differentiation operation is very complex and time-consuming, as the PGF function becomes a very high-order fraction for the cases with a large initial contention window size, maximum backoff stage, and retransmission limit.

To simplify analysis, we derive the first and second moments of the packet service time with the closed queueing network model in [14]. The WLAN channel is mapped to a queue server with state-dependent service rates, and multi-class customers with finite populations are served in a processor sharing (PS) discipline. With the contention-based MAC protocol in the WLAN, an admitted traffic flow is served packet by packet according to its contention with other flows. The service rates of the WLAN channel to sharing flows become dependent on the numbers of admitted flows. Packets from different ongoing flows take turns to be served, which is similar to time-sharing computer systems, where the research on PS queues originates from [12]. In time-sharing computer systems, the CPU computation power is time-slotted and a small quantum of one job is served each time. In such a PS manner, the CPU is shared by all jobs, and a job finishes until all of its pieces are served. In contrast, for the first come first served (FCFS) service discipline, the flow capturing the WLAN channel occupies the channel until all of its packets are served.

In the closed queueing network, each terminal is viewed as alternating between "think" mode (no service request) and "wait" (waiting for service completion) mode. For the $M/G/1$ queue of each flow, when the packet service rate provided by the WLAN channel for the specific flow is larger than the packet arrival rate on average (unsaturated case), the mean packet departure rate from the $M/G/1$ queue is equal to the mean packet arrival rate. The average think time, i.e., the mean time spent by a user in generating a new job, is approximated by the inverse of the mean packet departure rate. The packet service time is just the waiting time of the job in the closed queueing network.

To obtain the moments of the waiting time, the MAC contention and backoff process is analyzed to derive the state-dependent service rates of the WLAN channel. Given n_1 voice flows and n_2 data flows having backlogged packets in the corresponding queues, the probability that a packet from a voice flow collides with any other voice or data flow transmitting in the same slot is given by

$$p_v = 1 - (1-\tau_v)^{n_1 - 1}(1-\tau_d)^{n_2} \tag{6}$$

where τ_v and τ_d are the voice and data flow's transmission probabilities in a slot, respectively, given by [15]

$$\tau_v = \frac{\overline{R}_v}{\overline{R}_v + \overline{W}_v}, \qquad \tau_d = \frac{\overline{R}_d}{\overline{R}_d + \overline{W}_d} \tag{7}$$

in which $\overline{W}_v$ ($\overline{W}_d$) and $\overline{R}_v$ ($\overline{R}_d$) are the average backoff time and average number of transmission attempts of a target packet from a voice (data) flow, respectively. In particular,

$$\overline{W}_v = \frac{1}{2(1-p_v)(1-2p_v)}\Big[W(1-(2p_v)^{m'+1}) + W2^{m'}(p_v^{m'+1} - p_v^{m+1})(1-2p_v) - (1-2p_v)(1-p_v^{m+1})\Big]$$

$$\overline{R}_v = \sum_{i=0}^{m-1}\left[(i+1)\cdot p_v^i \cdot (1-p_v)\right] + (m+1)\cdot p_v^m = \frac{1-p_v^m}{1-p_v}$$

with $W = CW_{min} + 1$ and CW_{min} being the initial backoff window (e.g., it is 31 in IEEE 802.11), m the retransmission limit, and m' the maximum backoff stage. Similarly, we can obtain $\overline{W}_d$ and $\overline{R}_d$, and accordingly τ_d and the collision probability p_d of a data flow transmitting in a slot.

For data traffic with a relatively large packet payload size, the request to send (RTS)-clear to send (CTS) mechanism is considered to reduce the collision resolution time. The durations of a successful and a collided transmission from a data flow are respectively given by $T_{sd} = T_{DIFS} + T_{RTS} + T_{SIFS} + T_{CTS} + T_{SIFS} + T_{D\_DATA} + T_{SIFS} + T_{ACK}$ and $T_{cd} = T_{DIFS} + T_{RTS} + T_{CTS\_TO}$. Here, T_{D_DATA}, T_{RTS}, T_{CTS}, T_{ACK}, T_{DIFS}, T_{SIFS}, and T_{CTS_TO} are the transmission time of a DATA frame from a data flow, RTS frame duration, CTS frame duration, transmission time of an ACK frame, DCF interframe space (DIFS), short interframe space (SIFS), and waiting time for a CTS TIMEOUT, respectively. On the other hand, the time needed for a successful transmission from a voice flow is $T_{sv} = T_{DIFS} + T_{V\_DATA} + T_{SIFS} + T_{ACK}$, with T_{V_DATA} being the transmission time of a voice DATA frame. The time of a collided transmission from a voice flow (denoted by T_{cv}) depends on the traffic types of the collided frames. A target voice frame may collide with only voice frames with a probability $(1-\tau_d)^{n_2}\left[1-(1-\tau_v)^{n_1-1}\right]/p_v$. The target voice frame may also collide with at least one data frame with a probability $\left[1-(1-\tau_d)^{n_2}\right]/p_v$. We then have

$$T_{cv} = (T_{DIFS} + T_{V\_DATA} + T_{ACK\_TO}) \cdot (1-\tau_d)^{n_2} \left[1-(1-\tau_v)^{n_1-1}\right] \Big/ p_v + T_{cd} \cdot \left[1-(1-\tau_d)^{n_2}\right]/p_v$$

where T_{ACK_TO} is the waiting time for an ACK TIMEOUT. Let $\xi_1(n_1, n_2)$ and $\xi_2(n_1, n_2)$ denote the voice and data packet service rates, respectively. Based on an analytical method similar to those in [1, 15, 16], we further have

$$\left[\xi_1(n_1,n_2)\right]^{-1} = n_1 T_{sv} + n_2 T_{sd} + \overline{W}_v + \frac{1}{2}(n_1 \overline{T}_{cv} + n_2 \overline{T}_{cd})$$
$$\left[\xi_2(n_1,n_2)\right]^{-1} = n_1 T_{sv} + n_2 T_{sd} + \overline{W}_d + \frac{1}{2}(n_1 \overline{T}_{cv} + n_2 \overline{T}_{cd})$$

where $\overline{T}_{cv}$ and $\overline{T}_{cd}$ are the average collision time of a frame in a voice flow and that of a frame in a data flow, given by $\overline{T}_{cv} = \sum_{i=1}^{m} iT_{cv} \cdot p_v^i \cdot (1-p_v)$ and $\overline{T}_{cd} = \sum_{i=1}^{m} iT_{cd} \cdot p_d^i \cdot (1-p_d)$, respectively.

Suppose there are n_v^w voice flows and n_d^w data flows admitted in the WLAN. When $n_v^w \geq 1$, let $K_1 = n_v^w - 1$ and $K_2 = n_d^w$. First, we derive the equilibrium distribution of an equivalent birth-death process with two customer classes of finite populations K_1 and K_2 served in a PS discipline. The equilibrium distribution $\boldsymbol{\pi}$ can be obtained from the following linear equation system [14]

$$\boldsymbol{\pi}'\boldsymbol{F} = \boldsymbol{0}' \tag{8}$$

where $\boldsymbol{F}$ is defined as follows. Let $\boldsymbol{n} = (n_1, n_2)$. For $0 \leq n_1 \leq K_1, 0 \leq n_2 \leq K_2$, and $j = 1, 2$.

$$\begin{aligned} F(\boldsymbol{n}-\boldsymbol{e}_j, \boldsymbol{n}) &= \alpha_j(n_v^w, n_d^w) \cdot (K_j - n_j + 1) \\ F(\boldsymbol{n}+\boldsymbol{e}_j, \boldsymbol{n}) &= (n_j+1) \cdot \xi_j(\boldsymbol{n}+\boldsymbol{e}_j) \\ F(\boldsymbol{n}, \boldsymbol{n}) &= -\textstyle\sum_j \Big[\alpha_j(n_v^w, n_d^w) \cdot (K_j - n_j) \\ &\qquad + (1-\delta_{\boldsymbol{n},\boldsymbol{0}}) \cdot n_j \cdot \xi_j(\boldsymbol{n})\Big] \end{aligned} \tag{9}$$

where $\alpha_1(n_v^w, n_d^w) = \lambda_v^p$, $\alpha_2(n_v^w, n_d^w) = \lambda_d^p(n_v^w, n_d^w)$, $\boldsymbol{e}_j$ is a two-dimensional vector with unity in the jth element and zero elsewhere, and δ the Kronecker delta symbol. Then, the first and second moments of the voice packet service time $S_v(n_v^w, n_d^w)$ are given by [14]

$$\begin{aligned} \mathrm{E}[S_v(n_v^w, n_d^w)] &= -\boldsymbol{\pi}'\boldsymbol{M}^{-1}\boldsymbol{1} = \left[\xi_v^w(n_v^w, n_d^w)\right]^{-1} \\ \mathrm{E}[S_v^2(n_v^w, n_d^w)] &= -2\boldsymbol{\pi}'\boldsymbol{M}^{-2}\boldsymbol{1} \end{aligned} \tag{10}$$

where

$$\begin{aligned} M(\boldsymbol{n}-\boldsymbol{e}_j, \boldsymbol{n}) &= \alpha_j(n_v^w, n_d^w) \cdot (K_j - n_j + 1) \\ M(\boldsymbol{n}+\boldsymbol{e}_j, \boldsymbol{n}) &= (n_j+1) \cdot \xi_j(\boldsymbol{n}+\boldsymbol{e}_j+\boldsymbol{e}_J) \\ M(\boldsymbol{n}, \boldsymbol{n}) &= -\textstyle\sum_j \Big[\alpha_j(n_v^w, n_d^w) \cdot (K_j - n_j) \\ &\qquad + (n_j + \delta_{j,J}) \cdot \xi_j(\boldsymbol{n}+\boldsymbol{e}_J)\Big] \end{aligned} \tag{11}$$

and $\boldsymbol{1}$ is the vector with unity for all elements and $J = 1$ indicating that the derivation is for the packet service time of voice flows. Hence, when there are n_v^w voice flows and n_d^w data flows admitted in the WLAN, the average voice packet delay can be obtained by the P-K formula as

$$\mathrm{E}[T_v^p(n_v^w, n_d^w)] = \mathrm{E}[S_v(n_v^w, n_d^w)]\Big(1 + \frac{1+C_{v,v}^2}{2} \cdot \frac{\rho_v^p}{1-\rho_v^p}\Big)$$

where

$$C_{v,v}^2 = \frac{\mathrm{Var}[S_v(n_v^w, n_d^w)]}{\mathrm{E}^2[S_v(n_v^w, n_d^w)]}, \qquad \rho_v^p = \lambda_v^p \cdot \mathrm{E}[S_v(n_v^w, n_d^w)].$$

Similarly, when $n_d^w \geq 1$, the first and second moments of the data packet service time $S_d(n_v^w, n_d^w)$ can be obtained by setting $K_1 = n_v^w$, $K_2 = n_d^w - 1$, and $J = 2$. Then, $\mathrm{E}[S_d(n_v^w, n_d^w)] = [\xi_d^w(n_v^w, n_d^w)]^{-1}$.

D. Model for Cellular Capacity

In this study, we consider a code-division multiple access (CDMA) cellular system [17] with variable spreading gain. Voice traffic with adaptive multi-rate (AMR) codec at 12.2 Kbps is delivered with dedicated channels (DCH), while data traffic such as Web browsing sessions can be transported over the high-speed downlink shared channels (HS-DSCH). As the data services such as Web browsing is asymmetric and the downlink is more congested, the downlink cellular capacity is analyzed in the following. Since the common downlink bandwidth is shared by all users, the number of admitted users should be limited to bound the interference level and satisfy the QoS requirements for the ratio of bit energy to noise and interference power spectral density $\left(\frac{E_b}{N_0}\right)$. Let the cell capacity be represented by the maximum numbers of voice and data users that can be simultaneously admitted. Then, the downlink capacity can be evaluated based on a cell load factor [17], which is defined as

$$\eta_{DL} = \sum_{i=1}^{n_v^c} \frac{\rho + f_{DL}}{\frac{W_c}{\left(\frac{E_b}{N_0}\right)_v \beta_v R_{b,v}} + \rho} + \sum_{i=1}^{n_d^c} \frac{\rho + f_{DL}}{\frac{W_c}{\left(\frac{E_b}{N_0}\right)_d R_{b,d}} + \rho} \tag{12}$$

where n_v^c and n_d^c are the number of voice and data users, respectively, ρ is the orthogonality factor, f_{DL} is the ratio between intercell interference and total intracell power measured at the user receiver, W_c is the total cell bandwidth, $R_{b,v}$ ($R_{b,d}$) is the bit rate of voice (data) users, β_v the activity factor of voice users, and $\left(\frac{E_b}{N_0}\right)_v$ and $\left(\frac{E_b}{N_0}\right)_d$ are the $\frac{E_b}{N_0}$ requirements of voice and data users, respectively. Then, the power limitation of the BS is equivalent to bounding the cell load factor by

$$\eta_{max} = 1 - \frac{P_p + P_N X_n}{P_{T,max}} \tag{13}$$

where

$$X_n = \sum_{i=1}^{n_v^c} \frac{L_{p,i}}{\frac{W_c}{\left(\frac{E_b}{N_0}\right)_v \beta_v R_{b,v}} + \rho} + \sum_{i=1}^{n_d^c} \frac{L_{p,i}}{\frac{W_c}{\left(\frac{E_b}{N_0}\right)_d R_{b,d}} + \rho} \tag{14}$$

with $L_{p,i}$ being the path loss for the i^{th} user, P_p the power devoted to common control channels, P_N the background noise power, and $P_{T,max}$ the maximum transmitted power of the BS. From (12) - (14), we can derive the cell capacity region in terms of (n_v^c, n_d^c) vectors, in which the $\frac{E_b}{N_0}$ requirements of voice and data users are satisfied with the limited transmission power of the BS.

III. Admission Control for Cellular/WLAN Interworking

A. Distributed Admission Control

With the heterogeneous QoS support of the underlying integrated network, the voice and data traffic in a double-coverage area should be properly directed to the cell and WLAN. Due to the heterogeneity, especially when the two systems are loosely coupled, it is challenging for a central controller to obtain updated information of both systems (e.g., the numbers of ongoing calls in the integrated cell and WLANs) to make an optimal decision for each admission request. A high controlling overhead is induced since the signaling messages have to travel over a long path involving many network elements. To reduce the signaling overhead, the frequency of exchanging network information has to be reduced; but outdated network information may affect the decision accuracy. Consequently, in a loosely coupled cellular/WLAN network, distributed admission control is more practical, although the admission decision may not be optimal in terms of maximizing resource utilization with QoS guarantee.

Instead of applying complex criteria for network selection, the following simple admission scheme is proposed to investigate the dependence of resource utilization on traffic distribution and other important traffic and mobility parameters. An incoming voice (data) call in the double-coverage area requests admission to the cell with a probability θ_v^c (θ_d^c), while it requests admission to the WLAN with a probability $\theta_v^w = 1 - \theta_v^c$ ($\theta_d^w = 1 - \theta_d^c$). The admission parameters θ_v^c and θ_d^c (or θ_v^w and θ_d^w) are determined by the network for a given traffic load and broadcast to the associated MSs. Then, an MS can make a decision on its own according to the parameters θ_v^c and θ_d^c and send the admission request to the corresponding target network. With the simplicity, the proposed admission scheme can be implemented in a distributed manner. Also, as the network elements involved in the admission decision are limited to be as few as possible, the signaling overhead is reduced. The cellular network and the integrated WLANs only need to exchange information and update the above admission parameters with traffic variation. It is especially suitable for the cases where it is not affordable to make each admission decision based on the system states of both networks. However, as the proposed scheme distributes the incoming traffic load to the integrated cell and WLAN by properly controlling the assignment probabilities θ_v^c and θ_d^c, it achieves simplicity but may not fully exploit the performance gain achievable from the cellular/WLAN interworking.

Within the two-tier overlaying structure, the vertical handoff from the cell to the overlaying WLAN is not necessary but optional to maintain an ongoing call. Hence, the handoff traffic load to the WLAN can be controlled by properly adjusting the admission parameters of the WLAN, e.g., by using a simple guard channel policy. Let G_v^w (G_d^w) denote the number of guard channels reserved in the WLAN for new voice (data) calls from the double-coverage area. Then, a handoff voice (data) call from the cell is admitted to the WLAN if the number of voice (data) calls in the WLAN is less than $M_v^w = N_v^w - G_v^w$ ($M_d^w = N_d^w - G_d^w$) and rejected otherwise.

The admission parameters θ_v^w and θ_d^w (or $\theta_v^c = 1 - \theta_v^w$ and $\theta_d^c = 1 - \theta_d^w$) are selected so as to maximize the acceptable traffic load with a given cell/WLAN cluster, which implies a high utilization. This admission strategy can be extended and applied in combination with other network selection criteria. For example, the cellular network also differs from WLANs in pricing rates. The service cost in the WLAN is generally much lower than that in the cellular network. A mobile user may prefer to get admitted to the WLAN for the low cost or to the cellular network if the service quality is more important. Suppose that an incoming call requests admission to the WLAN with a probability γ^w, while it requests admission to the cell with a probability $\gamma^c = 1 - \gamma^w$. Let $\omega_i,\ i = 1, 2, ..., r$, denote the relative weights of r different criteria and θ_i^w (θ_i^c) the probability of selecting the WLAN (cell) as the target when the i^{th} criterion is considered. Then, we have

$$\gamma^w = \sum_{i=1}^{r} \omega_i \cdot \theta_i^w, \qquad \gamma^c = \sum_{i=1}^{r} \omega_i \cdot \theta_i^c \tag{15}$$

where

$$\sum_{i=1}^{r} \omega_i = 1, \quad \theta_i^w = 1 - \theta_i^c, \quad i = 1, 2, ..., r. \tag{16}$$

Considering both the overall utilization maximization and user preference, the probabilities of selecting the WLAN and cell as the admission target are respectively given by

$$\gamma^w = \omega_1 \cdot \theta_1^w + \omega_2 \cdot \theta_2^w, \ \ \gamma^c = \omega_1 \cdot \theta_1^c + \omega_2 \cdot \theta_2^c = 1 - \gamma_w \tag{17}$$

where $\theta_1^w = \theta_d^w$ and $\theta_1^c = \theta_d^c$ for an incoming data call; $\theta_1^w = \theta_v^w$ and $\theta_1^c = \theta_v^c$ for a voice call; θ_2^w and θ_2^c can be configured according to whether the user prefers the low cost and high data speed of the WLAN or the guaranteed real-time service quality of the cell; the weights ω_1 and ω_2 can be configured based on the relative importance of the two criteria.

B. QoS Evaluation for Voice and Data Services in the WLAN

By configuring the admission parameters θ_v^c and θ_d^c (or θ_v^w and θ_d^w) of the distributed admission scheme, the voice and data traffic load is properly distributed to the overlaying cell and WLAN. Let λ_{v1} and λ_{v2} (λ_{d1} and λ_{d2}) denote the mean arrival rates of new voice (data) calls in the cellular-only area and double-coverage area, respectively. Then, the mean arrival rates of new voice and data calls to the cell from the double-coverage area are $\lambda_{nv2}^c = \theta_v^c \cdot \lambda_{v2}$ and $\lambda_{nd2}^c = \theta_d^c \cdot \lambda_{d2}$, respectively. Similarly, the mean arrival rates of new voice and data calls to the WLAN from the double-coverage area are $\lambda_{nv}^w = \theta_v^w \cdot \lambda_{v2}$ and $\lambda_{nd}^w = \theta_d^w \cdot \lambda_{d2}$, respectively. To determine the admission parameters, the QoS metrics of interest need to be evaluated accurately and effectively. We begin with the WLAN QoS.

The voice traffic load offered to the WLAN includes new calls from the double-coverage area with a mean rate λ_{nv}^w and handoff calls from the overlaying cell with a mean rate λ_{hv}^{c-w}. With Poisson arrivals and exponential call duration, the voice calls in the WLAN can be modeled by an $M/G/K/K$ queue, where the channel occupancy time $\min(T_v, T_r^w)$ is not exponential. With the insensitive property of $M/G/K/K$ queues [18], the steady-state probability of voice calls in the WLAN is given by

$$\pi_v^w(j) = \pi_v^w(0) \prod_{l=1}^{j} \frac{\lambda_v^w(l)}{l \cdot \mu_v^w}, \quad j = 1, \dots, N_v^w \tag{18}$$

where

$$\frac{1}{\mu_v^w} = \mathrm{E}[\min(T_v, T_r^w)] = \frac{a}{a+1} \cdot \frac{1}{a\eta^w + \mu_v} + \frac{1}{a+1} \cdot \frac{1}{\frac{1}{a}\eta^w + \mu_v}$$

$\lambda_v^w(l) = \lambda_{nv}^w + \lambda_{hv}^{c-w}$ when $l \le M_v^w$, and $\lambda_v^w(l) = \lambda_{nv}^w$ when $M_v^w + 1 \le l \le N_v^w$. The voice call blocking probability of the WLAN is then $B_v^w = \pi_v^w(N_v^w)$, and the rejection probability for handoff voice calls from the cell is $D_v^w = \sum_{i=M_v^w}^{N_v^w} \pi_v^w(i)$.

The data call blocking probability and rejection probability of the WLAN can be evaluated similarly except for a state-dependent service rate. Given that the data call size follows the hyper-exponential distribution in (5), the data traffic to the WLAN can be viewed as two virtual classes [19] with Poisson arrival rates $\frac{b}{b+1}\lambda_d^w(j)$ and $\frac{1}{b+1}\lambda_d^w(j)$, respectively, and exponentially distributed service requirements with mean f_d/b and $b \cdot f_d$, respectively, where $\lambda_d^w(j) = \lambda_{nd}^w + \lambda_{hd}^{c-w}$ when $j \le M_d^w$, and $\lambda_d^w(j) = \lambda_{nd}^w$ when $M_d^w + 1 \le j \le N_d^w$. With i voice calls and j data calls admitted in the WLAN, the average channel occupancy time of data calls of the two virtual classes is respectively

$$\frac{1}{\mu_{d1}^w(i,j)} = \frac{a}{a+1} \cdot \frac{1}{a\eta^w + \frac{\chi_d^w(i,j)}{f_d/b}} + \frac{1}{a+1} \cdot \frac{1}{\frac{1}{a}\eta^w + \frac{\chi_d^w(i,j)}{f_d/b}}$$

$$\frac{1}{\mu_{d2}^w(i,j)} = \frac{a}{a+1} \cdot \frac{1}{a\eta^w + \frac{\chi_d^w(i,j)}{b \cdot f_d}} + \frac{1}{a+1} \cdot \frac{1}{\frac{1}{a}\eta^w + \frac{\chi_d^w(i,j)}{b \cdot f_d}}$$

where $\chi_d^w(\cdot)$ is the average transmission rate (in bps) for packets from a data call. Then, the steady-state probability of data calls in the WLAN is approximated by $\pi_d^w(j) = \sum_{i=0}^{N_v^w} \pi_v^w(i)\tilde{\pi}_d^w(j|i)$, where

$$\tilde{\pi}_d^w(j|i) = \tilde{\pi}_d^w(0|i) \prod_{l=1}^{j} \left[\frac{\frac{b}{b+1}\lambda_d^w(l)}{l \cdot \mu_{d1}^w(i,l)} + \frac{\frac{1}{b+1}\lambda_d^w(l)}{l \cdot \mu_{d2}^w(i,l)} \right]$$
$$j = 1, \dots, N_d^w. \tag{19}$$

Similarly, the data call blocking probability of the WLAN is obtained as $B_d^w = \pi_d^w(N_d^w)$, and the rejection probability for handoff data calls from the cell is $D_d^w = \sum_{i=M_d^w}^{N_d^w} \pi_d^w(i)$. From the Little's law, the average throughput of each admitted data call in the WLAN is given by

$$\Gamma_d^w = f_d \left[\sum_{i=0}^{N_v^w} \pi_v^w(i) \frac{\sum_{l=1}^{N_d^w} l \cdot \tilde{\pi}_d^w(l|i)}{\lambda_{nd}^w \cdot (1 - B_d^w) + \lambda_{hd}^{c-w} \cdot (1 - D_d^w)} \right]^{-1}. \tag{20}$$

C. QoS Evaluation for Voice and Data Services in the Cell

The QoS performance experienced by voice and data calls in the cell is more complex to evaluate than that in the WLAN. For example, the location-dependent mobility in a cell results in different channel occupancy time for traffic in the cellular-only area and double-coverage area. This necessitates differentiation of traffic from different areas in the QoS evaluation. To effectively evaluate the QoS metrics while searching for the admission parameters, we circumvent the computation complexity of solving large-scale balance equations [3] by using MGFs. Depending on the WLAN state, the average channel occupancy time of voice calls in the cellular-only area can be derived from (2) as

$$\mathrm{E}[\min(T_v, T_{r1}^c)] = \frac{1}{\mu_v} - \frac{1}{\mu_v}\Phi_1(-\mu_v) \tag{21}$$
$$= \frac{1}{\mu_v} - \frac{1}{\mu_v} p^{c-c} \frac{\eta^c}{\eta^c + \mu_v} \frac{1}{1 - p^{c-w}\psi(-\mu_v)} \triangleq \frac{1}{\mu_{v1}^c}$$

when there is not enough free capacity in the WLAN for a voice call; and it is $1/(\mu_v + \eta^c)$ when the incoming voice call can be admitted to the WLAN. Similarly, for voice calls in the double-coverage area, the average channel occupancy time is $\frac{1}{\mu_v} - \frac{1}{\mu_v}\psi(-\mu_v)$ if there is free room for one more voice call in the WLAN or obtained from (3) as

$$\mathrm{E}[\min(T_v, T_{r2}^c)] = \frac{1}{\mu_v} - \frac{1}{\mu_v}\Phi_2(-\mu_v) \tag{22}$$
$$= \frac{1}{\mu_v} - \frac{1}{\mu_v} p^{c-c} \frac{\psi(-\mu_v)}{1 - p^{c-w}\psi(-\mu_v)} \triangleq \frac{1}{\mu_{v2}^c}$$

otherwise. To simplify analysis, we take an average for the mean service rates of voice calls in the cellular-only area and double-coverage area, which are respectively given by

$$\tilde{\mu}_{v1}^c = D_v^w \mu_{v1}^c + (1 - D_v^w) \cdot (\mu_v + \eta^c) \tag{23}$$

$$\tilde{\mu}_{v2}^c = D_v^w \mu_{v2}^c + (1 - D_v^w) \frac{\mu_v}{1 - \psi(-\mu_v)}. \tag{24}$$

By modeling the voice traffic in the cell with a multi-service loss system, we can obtain the steady-state probability of voice calls in the cell as

$$\pi_v^c(j) = \pi_v^c(0) \left(\frac{\lambda_{v1} + \lambda_{hv}^{c-c} + \lambda_{hv}^{w-c}}{\tilde{\mu}_{v1}^c} + \frac{\lambda_{nv2}^c}{\tilde{\mu}_{v2}^c} \right)^j \Big/ j!$$
$$j = 1, \dots, N_v^c \tag{25}$$

where λ_{hv}^{c-c} and λ_{hv}^{w-c} are mean arrival rates of handoff voice calls from neighboring cells and the overlaying WLAN, respectively. Then, the voice call blocking/dropping probability of the cell is $B_v^c = \pi_v^c(N_v^c)$. New call blocking and handoff call dropping are not differentiated in the cell for simplicity. In [4], we have studied the differentiation of new calls, horizontal handoff calls, and vertical handoff calls with a randomized guard channel policy.

Two other important aspects need to be properly addressed to evaluate the QoS performance of data calls in the cell. First, with the restricted access mechanism, the data call service rates become dependent on both active voice and data calls in the cell, as all bandwidth unused by current voice traffic is shared equally by active data calls. Second, the high variability of data call size as shown in (5) should be properly dealt with in the QoS evaluation. Under the assumption that the number of data calls fluctuates in a much smaller time scale than that of voice calls, the analysis for data traffic can be approximately decomposed from that for voice. As such, for data traffic, the cell can be modeled by a symmetric queue [18] serving multiple classes. Similar to the QoS evaluation of data in the WLAN, both data calls admitted into the cell from the cellular-only area and those from the double-coverage area are respectively differentiated into two virtual classes.

Given j voice calls and k data calls in the cell, similar to (23) and (24), the service rates of the two virtual classes of data calls in the cellular-only area can be approximated by

$$\tilde{\mu}_{d1}^{c1}(j,k) = D_d^w \mu_{d1}^{c1}(j,k) + (1-D_d^w)\left[b\cdot\nu_d^c(j,k) + \eta^c\right]$$
$$\tilde{\mu}_{d2}^{c1}(j,k) = D_d^w \mu_{d2}^{c1}(j,k) + (1-D_d^w)\left[\nu_d^c(j,k)/b + \eta^c\right]$$

where

$$\nu_d^c(j,k) = \frac{R_{b,d}}{f_d}, \qquad \mu_{d1}^{c1}(j,k) = \frac{b\cdot\nu_d^c(j,k)}{1-\Phi_1(-b\cdot\nu_d^c(j,k))}$$
$$\mu_{d2}^{c1}(j,k) = \frac{\nu_d^c(j,k)/b}{1-\Phi_1(-\nu_d^c(j,k)/b)}.$$

Similarly, the service rates of the two virtual classes of data calls admitted to the cell from the double-coverage area can be obtained as

$$\tilde{\mu}_{d1}^{c2}(j,k) = D_d^w \mu_{d1}^{c2}(j,k) + (1-D_d^w)\frac{b\cdot\nu_d^c(j,k)}{1-\psi(-b\cdot\nu_d^c(j,k))}$$
$$\tilde{\mu}_{d2}^{c2}(j,k) = D_d^w \mu_{d2}^{c2}(j,k) + (1-D_d^w)\frac{\nu_d^c(j,k)/b}{1-\psi(-\nu_d^c(j,k)/b)}$$

where

$$\mu_{d1}^{c2}(j,k) = \frac{b\cdot\nu_d^c(j,k)}{1-\Phi_2(-b\cdot\nu_d^c(j,k))}$$
$$\mu_{d2}^{c2}(j,k) = \frac{\nu_d^c(j,k)/b}{1-\Phi_2(\nu_d^c(j,k)/b)}.$$

Hence, given that there are i voice calls in the cell, the equilibrium distribution of the symmetric queue for data traffic in the cell is given by

$$\tilde{\pi}_d^c(j|i) = \tilde{\pi}_d^c(0|i)\prod_{l=1}^{j}\left[\frac{\frac{b}{b+1}\lambda_{d1}^c}{l\cdot\tilde{\mu}_{d1}^{c1}(i,l)} + \frac{\frac{1}{b+1}\lambda_{d1}^c}{l\cdot\tilde{\mu}_{d2}^{c1}(i,l)} + \frac{\frac{b}{b+1}\lambda_{nd2}^c}{l\cdot\tilde{\mu}_{d1}^{c2}(i,l)} + \frac{\frac{1}{b+1}\lambda_{nd2}^c}{l\cdot\tilde{\mu}_{d2}^{c2}(i,l)}\right] \quad (26)$$

TABLE II
EVALUATION PROCEDURE FOR ADMISSION PARAMETERS

Derive the cellular capacity in terms of vectors (n_v^c, n_d^c) from (12) - (14) to meet the $\frac{E_b}{N_0}$ requirements
Derive the WLAN capacity in terms of vectors (n_v^w, n_d^w) from (6) - (10) to meet the stability constraints
$N_{v,max}^w = \max(n_v^w)$, $N_{v,max}^c = \max(n_v^c)$
for $N_v^w = 0$ to $N_{v,max}^w$ **do** // Evaluation for voice traffic
  for $G_v^w = 0$ to N_v^w **do**
    Derive the maximum θ_v^w from (18) so that $B_v^w \le Q_{vb}$
    Record the corresponding steady-state probabilities of voice calls in the WLAN $\pi_v^w(i)$, $i = 0, 1, ..., N_v^w$
    Derive the minimum N_v^c $(\le N_{v,max}^c)$ from (25) so that $B_v^c \le Q_{vb}$
    Record the corresponding steady-state probabilities of voice calls in the cell $\pi_v^c(i)$, $i = 0, 1, ..., N_v^c$
  end for
end for
for $N_v^w = 0$ to $N_{v,max}^w$ **do** // Evaluation for data traffic
  $N_{d,max}^w = \max(n_d^w)$ with $n_v^w = N_v^w$
  for $G_v^w = 0$ to N_v^w **do**
    for $N_d^w = 0$ to $N_{d,max}^w$ **do**
      for $G_d^w = 0$ to N_d^w **do**
        Initialize $\lambda_{d,min}$ and $\lambda_{d,max}$
        $\lambda_d \leftarrow (\lambda_{d,min} + \lambda_{d,max})/2$
        loop
          Derive the maximum θ_d^w from (19) and (20) so that $B_d^w \le Q_{db}$ and $\Gamma_d^w \ge Q_{dt}$
          Derive the minimum N_d^c from (26) and (27) so that $B_d^c \le Q_{db}$ and $\Gamma_d^c \ge Q_{dt}$
          if Solutions for θ_d^w and N_d^c exist **then** // The given traffic load λ_d is acceptable
            $\lambda_{d,min} \leftarrow \lambda_d$
            $\lambda_d \leftarrow (\lambda_{d,min} + \lambda_{d,max})/2$
          else
            $\lambda_{d,max} \leftarrow \lambda_d$
            $\lambda_d \leftarrow (\lambda_{d,min} + \lambda_{d,max})/2$
          end if
          if The acceptable λ_d converges **then**
            Exit loop
          end if
        end loop
        Record the maximum acceptable λ_d
      end for
    end for
  end for
end for

TABLE III
SYSTEM PARAMETERS

Parameter	Value	Parameter	Value
$(\eta^c)^{-1}$	10 min	$(\eta^w)^{-1}$	14 min
p^{c-c}	0.76	p^{c-w}	0.24
W_c	3.84 Mchips/s	$R_{b,v}$	12.2 Kbps
$(\mu_v)^{-1}$	140 s	f_d	64 Kbytes
Q_{vb} (Q_{db})	0.01	Q_{dt}	200 Kbps
m'	5	m	7
CW_{min}	7, 31	Slot	20 μs
T_{SIFS}	10 μs	T_{DIFS}	50 μs
T_{RTS}	13.6 slots	T_{CTS}	12.4 slots
T_{ACK}	10.2 slots	λ_v^p	0.0004 frames/slot
Voice payload	50 bytes/packet	Data payload	1000 bytes/packet
ρ	0.4	f_{DL}	0.55
β_v	0.43	$P_{T,max}$	43 dB
P_p	33 dB	P_N	-106 dB
$\left(\frac{E_b}{N_0}\right)_v$	4.57 dB	$\left(\frac{E_b}{N_0}\right)_d$	4.69 dB

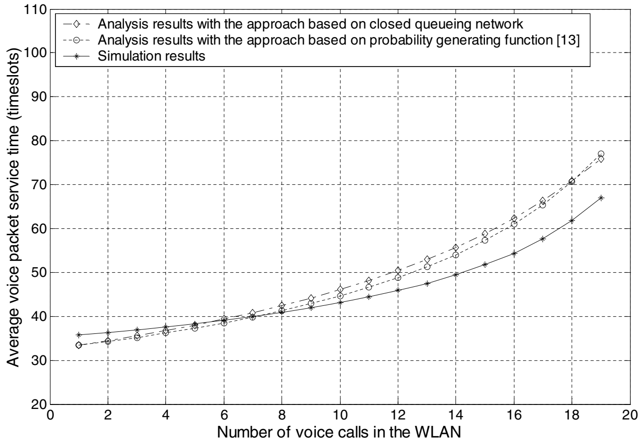

Fig. 2. Average voice packet service time in the WLAN ($\mathrm{E}[S_v]$).

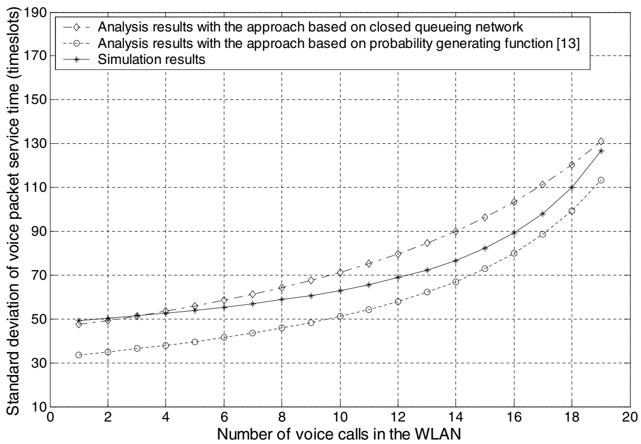

Fig. 3. Standard deviation of voice packet service time in the WLAN ($\sqrt{\mathrm{E}[S_v^2] - \mathrm{E}^2[S_v]}$).

$$\lambda_{d1}^c = \lambda_{d1} + \lambda_{hd}^{c-c} + \lambda_{hd}^{w-c}, \quad i = 0, ..., N_v^c, \quad j = 1, ..., N_d^c$$

where λ_{d1}^c is the total mean arrival rate of data calls in the cellular-only area, λ_{hd}^{c-c} and λ_{hd}^{w-c} are the mean arrival rates of handoff data calls from neighboring cells and the overlaying WLAN, respectively. Let $\pi_d^c(\cdot)$ denote the steady-state probability of data calls in the cell. Then, $\pi_d^c(j) = \sum_{i=0}^{N_v^c} \pi_v^c(i)\tilde{\pi}_d^c(j|i)$, $j = 0, 1, ..., N_d^c$, and the data call blocking/dropping probability of the cell is $B_d^c = \pi_d^c(N_d^c)$. From the Little's law, the average data call throughput in the cell can then be obtained as

$$\Gamma_d^c = f_d \left[\sum_{i=0}^{N_v^c} \pi_v^c(i) \frac{\sum_{l=1}^{N_d^c} l \cdot \tilde{\pi}_d^c(l|i)}{(\lambda_{d1}^c + \lambda_{nd2}^c) \cdot (1 - B_d^c)} \right]^{-1}. \quad (27)$$

In the preceding analysis, the call blocking/dropping probabilities and data call throughput can be derived from the steady-state probabilities of voice and data calls in the cell and WLAN, which are dependent on the mean arrival rates of handoff calls. With the statistical equilibrium assumption, the mean arrival rates of handoff calls are in turn dependent on call blocking/dropping probabilities. This inter-dependence requires that the QoS metrics be evaluated recursively.

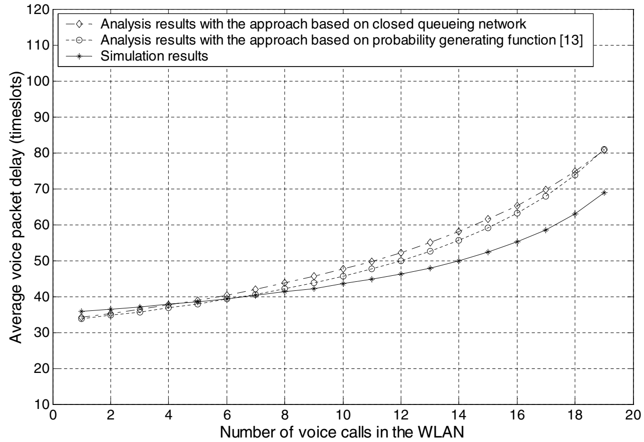

Fig. 4. Average voice packet delay in the WLAN ($\mathrm{E}[T_v^p]$).

IV. Numerical Results and Discussion

By applying the QoS evaluation given in Section III in a search algorithm similar to that given in [1], we can obtain the best configuration for admission parameters θ_v^w and θ_d^w to maximize the admissible traffic load with the given cell/WLAN cluster. For a target traffic load, the corresponding voice and data admission regions of the cell and WLAN $(N_v^c, N_d^c, N_v^w, G_v^w, N_d^w, G_d^w)$ can be determined with the admission parameters θ_v^w and θ_d^w. Table II summarizes the evaluation procedure. For time-varying traffic, the load fluctuation should be kept track of by periodical measurement and estimation. As the proposed QoS evaluation is performed with close-form approximations, it is reasonable to adaptively adjust the admission parameters for a target traffic load. In this section, we validate our analytical approach and investigate the impact of admission parameters on resource utilization.

A. Validation of WLAN Model for Capacity Analysis

In Section II-C, we analyze the packet service time with a closed queueing network model and derive the WLAN capacity accordingly. To verify the accuracy of our analysis, we compare our analysis results with computer simulation results. Since quite a few bugs are found in the WLAN MAC implementation of ns-2 simulator [20], we develop our own event-driven simulator with C/C++. Our analysis and simulation results are further compared with those obtained with the approach in [13] based on probability generating functions. Given in Table III are the system parameters, which are selected by referring to the popular WLAN standard IEEE 802.11 DCF [21].

Figs. 2-4 show the analysis and simulation results of the mean and standard deviation of voice packet service time (S_v) and average voice packet delay (T_v^p), respectively. Here, the initial contention window size is chosen to be 7. It is observed that the relative error between the analysis results and the simulation results is bounded to be less than 10% in most unsaturated cases. The overestimation is attributed to the following reasons. First, the interval between packet departures from each flow's service queue is assumed to be exponential for simplicity, which is not the real case. Moreover, the traffic flow is viewed as a fluid in the queueing analysis. In practical WLAN access, the traffic is serviced on

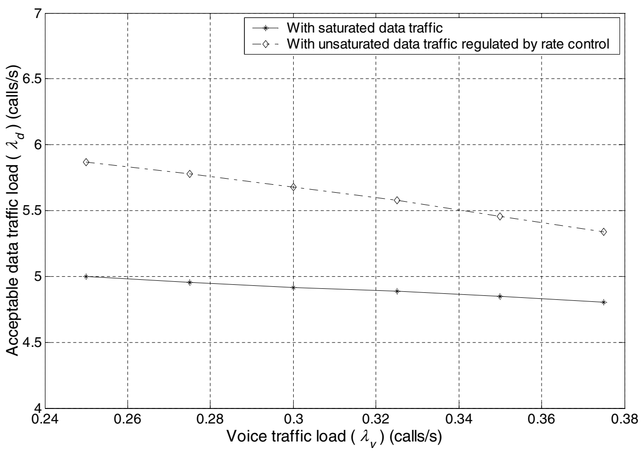

Fig. 5. Acceptable data traffic load (λ_d) with and without rate control to data traffic.

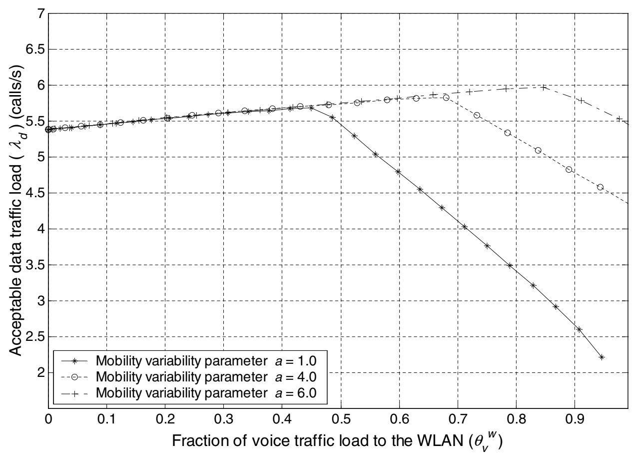

Fig. 6. The acceptable data traffic load (λ_d) versus the fraction of voice traffic carried by the WLAN (θ_v^w) with different mobility parameters (a).

a packet-by-packet basis. The coarse granularity of packets also accounts for the gap between the simulation results and analysis results.

It is observed that our analysis results match well with those of the approach in [13]. However, the approach in [13] involves differentiation of high-order fractions, which is very complex and time-consuming, especially for the cases with a large initial contention window size (CW_{min}), maximum backoff stage, and retransmission limit. When $CW_{min} \geq 31$, no valid results for the second moment of the packet service time can be obtained by Maple 9.5 [22]. In contrast, our analysis is based on linear computation and the evaluation can be executed very fast with a negligible error. Another important observation is that the average voice packet delay in the unsaturated cases is less than 20 ms, which is sufficiently small to meet the delay requirement of real-time voice service.

B. Accuracy of the Call-level QoS Evaluation Approach

In this study, we extend the MGF-based QoS evaluation approach proposed in [4]. As the QoS provisioning to user traffic admitted to the cell is more difficult to evaluate due to the location-dependent user mobility and vertical handoff, we compare the analysis and simulation results for the cell QoS in terms of voice call blocking/dropping probability (B_v^c), data call blocking/dropping probability (B_d^c), and mean data call throughput (Γ_d^c). Due to space limitation, the results are not shown here. It is observed that the analysis results agree well with the simulation results. The average relative error of B_v^c and Γ_d^c is less than 2%. Due to the elasticity and bandwidth adaptation of data calls, it is complicate to model statistically the data call duration. As a result, the average relative error of B_d^c is higher than that of B_v^c but still less than 8%. Hence, using this MGF-based approach, we can evaluate the QoS metrics effectively and accurately to determine the admission parameters.

C. Improvement of Resource Utilization via Rate Control to Data Traffic

Given a target voice traffic load ($\lambda_v = \lambda_{v1} + \lambda_{v2}$), the acceptable data traffic load ($\lambda_d = \lambda_{d1} + \lambda_{d2}$) can be maximized by properly selecting the admission parameters such as θ_v^w and θ_d^w. As discussed in Section II-C, rate control is considered for data traffic instead of allowing saturated data traffic as in [1]. As an example, Fig. 5 shows the improvement of the data traffic load λ_d that can be carried with QoS guarantee when rate control is applied to data traffic. When there is no rate control applied to the greedy data traffic, the number of data calls admitted to the WLAN has to be very restricted so that the voice delay requirement is not violated due to a high collision probability. Although a high throughput is provided to admitted data calls, not many data calls in the double-coverage area can be admitted to the WLAN. The data traffic load in the relatively congested cell cannot be effectively reduced. On the contrary, by applying rate control, more data calls can be admitted into the WLAN with a reasonable throughput and then effectively balance the data traffic load from the cell; at the same time, the voice delay requirement is guaranteed by keeping the collision probability small enough with traffic regulation.

D. Dependence of Utilization on Admission Parameters

Fig. 6 shows the dependence of the acceptable data traffic load (λ_d) on the admission parameter θ_v^w, which is the probability that an incoming voice call in the double-coverage area requests admission to the WLAN. The results are obtained with $\lambda_{v1} = 0.1$ calls/s, $\lambda_{v2} = 0.2$ calls/s, and $G_v^w = N_v^w$, i.e., no voice handoff from the cell to the overlaying WLAN. It can be seen that there exist optimal values of θ_v^w that maximize the acceptable traffic load, resulting in a maximum utilization. Similar phenomenon is observed in [1]. It is actually a result of properly balancing the voice and data traffic load in the cell and WLAN. On one hand, when more voice traffic in the double-coverage area is directed to the WLAN (i.e., a larger θ_v^w), more cellular bandwidth is available to admit data calls. With a smaller bandwidth, the cell is actually the bottleneck of the whole integrated system for data traffic. Hence, with the load balancing of the WLAN, the congestion of the cell and in turn the whole system is effectively relieved. On the other hand, with a larger θ_v^w, the maximum number of voice calls allowed in the WLAN (N_v^w) should also be larger to meet the voice call blocking/dropping probability requirement. However, the WLAN is very inefficient in supporting voice

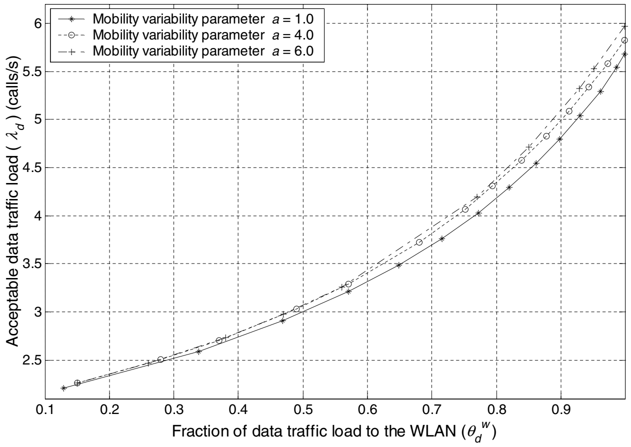

Fig. 7. The acceptable data traffic load (λ_d) versus the fraction of data traffic carried by the WLAN (θ_d^w) with different mobility parameters (a).

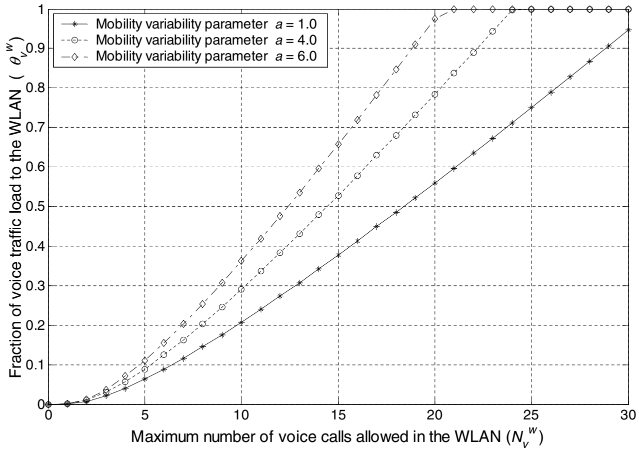

Fig. 8. The fraction of voice traffic carried by the WLAN (θ_v^w) versus the maximum number of voice calls allowed in the WLAN (N_v^w) with different mobility parameters (a).

traffic. Due to the small coverage of the WLAN, not only can frequent vertical handoffs between the cell and WLAN degrade the voice quality and increase the risk of call dropping, but also it is detrimental to multiplexing gain by breaking a call into more service stages. Although the voice traffic load to the cell is reduced to an extent, the number of data calls that can be accommodated by the WLAN is also significantly reduced. As a result, the total acceptable traffic load starts to decrease when θ_v^w is larger than a threshold. Hence, the admission parameter θ_v^w and correspondingly N_v^w should be properly determined to maximize the utilization.

Fig. 7 shows the variation of the acceptable data traffic load (λ_d) with θ_d^w, i.e., the probability that a data call in the double-coverage area requests admission to the WLAN, which is correlated with θ_v^w. With a smaller θ_v^w and less voice traffic load carried by the WLAN, θ_d^w can be larger to admit more data traffic and provide enough bandwidth for each admitted voice and data call in the WLAN. In Fig. 6, before θ_v^w reaches the point for a maximum acceptable data traffic load, θ_v^w is less than 0.85 and N_v^w less than 20 is enough to meet the bound for voice call blocking/dropping probability. For these cases, a high data call throughput is achievable since the number of voice calls allowed in the WLAN is quite restricted, and θ_d^w can be as large as 99% to carry almost all the data traffic in the double-coverage area. Within this region, a larger θ_v^w alleviates more voice traffic from the cell but does not affect much the data service in the WLAN, which results in a larger acceptable λ_d. On the other hand, when θ_v^w and corresponding N_v^w are further increased, θ_d^w is even smaller and the WLAN cannot carry a large portion of the data traffic load in the double-coverage area. The bottleneck effect of the cell becomes evident and the acceptable λ_d decreases, which is clearly shown in Fig. 7. In conclusion, with a small θ_d^w and a very large θ_v^w, the effectiveness of the WLAN to complement the cell is greatly jeopardized. The acceptable traffic load can be maximized by properly controlling θ_v^w and selecting $\theta_d^w \geq 90\%$. The θ_v^w should be large enough to balance the voice traffic load from the cell and also small enough to avoid the inefficient states of the WLAN in voice support.

By comparing the curves in Fig. 6 and Fig. 7, it is observed that, with a larger parameter a, the acceptable data traffic

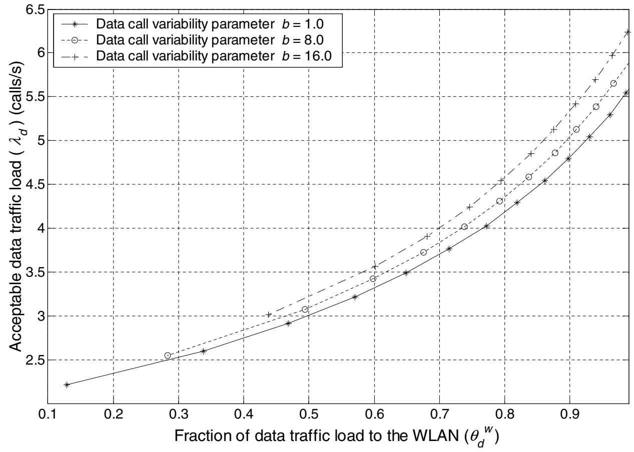

Fig. 9. The acceptable data traffic load (λ_d) versus the fraction of data traffic carried by the WLAN (θ_d^w) with different data call variability parameters (b).

load (λ_d) is larger. That is, a higher utilization is achievable when the variability of user mobility in the double-coverage area is higher. As illustrated in Fig. 6, when a is 1.0, 4.0, and 6.0, the highest utilization is achieved with different values of θ_v^w, which are 0.44, 0.68, and 0.84, respectively. A larger parameter a indicates that more users staying within the WLAN for a shorter time. As seen from Fig. 8, given a fixed N_v^w (i.e., the maximum number of voice calls allowed in the WLAN), when the parameter a is larger, a larger fraction of the voice calls in the double-coverage area can be carried by the WLAN and relieved from the cell. Therefore, the data call throughput in the cell is higher and more traffic is acceptable with QoS guarantee.

The data call variability also affects the admission parameters and resource utilization. As shown in Fig. 9, more data traffic is acceptable with a larger value of b, which indicates a higher variability of the data call size. With a fixed mean value, when the variability of the data call size is higher, more data calls have a smaller size and less have an extremely larger size. Hence, more data calls have a shorter channel occupancy time and can be carried by the WLAN with a high throughput. Also, more data calls can finish within the WLAN instead of having to hand over to the cell when users move out of the WLAN. The traffic load of the small-bandwidth cell is then effectively reduced, which results in a higher utilization. From Figs. 6-9, we can conclude that the first-order mean values themselves are not enough to determine the admission parameters. With the MGF-based approach, we can take into account higher-order statistics with reasonable complexity.

As discussed in Section III, the admission parameters G_v^w and G_d^w can be used to adjust the amounts of new and vertical handoff traffic admitted in the WLAN. It is found from the numerical results that the acceptable traffic load varies negligibly with G_d^w for a given N_d^w. With a large bandwidth and high data call throughput in the WLAN, the handoff traffic is much less than the new traffic and cannot significantly affect its load condition. Similarly, the acceptable traffic load (also the utilization) is only slightly dependent on G_v^w. This is because the number of voice calls admitted into the WLAN is quite limited, which also results in much less voice handoff traffic than new traffic.

Although the specific admission schemes in [1, 2] are different from the one in study, both can maximize the utilization by properly adjusting the admission parameters of the corresponding schemes to reach the best traffic distribution. For example, the service-differentiated admission scheme in [2] is equivalent to setting $G_v^w = N_v^w$ in the general admission scheme of this study. Then, to achieve the optimal θ_v^w and maximize the utilization, the maximum number of voice calls admitted to the cell from the double-coverage area (M_v^c) can be selected to be larger (smaller), so that more (less) voice traffic is directed to the cell to achieve a smaller (larger) θ_v^w. On the other hand, the WLAN-first scheme [1] is actually a special case of the general admission scheme with $G_v^w = 0$. In the WLAN-first scheme, to admit more (less) voice traffic in the double-coverage area to the cell and achieve a smaller (larger) θ_v^w, a smaller (larger) value can be chosen for the parameter G_{v2}^c, which represents the guard channels reserved for new and handoff voice traffic in the cellular-only area.

V. Conclusion

In this paper, we study how to improve voice and data service support in the cellular/WLAN integrated network by applying admission control. To take advantage of the complementary QoS provisioning capabilities of the two networks, the voice and data traffic should be properly admitted to the cell and WLAN by configuring the corresponding admission parameters. First, the WLAN is properly modeled to analyze the capacity. The voice packet delay can be bounded by restricting the numbers of voice and data calls admitted in the WLAN and applying rate control to regulate data traffic. Second, a general admission scheme is analyzed to investigate the dependence of resource utilization on admission parameters and traffic distribution. In order to determine the admission parameters, the analytical approach based on MGFs is extended to effectively and accurately evaluate the QoS metrics such as call blocking/dropping probabilities and data call throughput. It is observed from the numerical results that the resource utilization can be maximized when a balance is achieved in distributing the voice and data traffic load to the overlaying cell and WLAN.

References

[1] W. Song, H. Jiang, and W. Zhuang, "Call admission control for integrated voice/data services in cellular/WLAN interworking," in *Proc. IEEE ICC'06*, vol. 12, June 2006, pp. 5480–5485.

[2] W. Song and W. Zhuang, "QoS provisioning via admission control in cellular/wireless LAN interworking," in *Proc. 2nd Int'l. Conf. on Broadband Networks (BROADNETS'05)*, vol. 1, Oct. 2005, pp. 585–592.

[3] W. Song and W. Zhuang, "Admission control in cellular/WLAN integrated networks with unique mobility and traffic characteristics," submitted to *IEEE Trans. Wireless Commun.*

[4] W. Song, H. Jiang, and W. Zhuang, "Performance analysis of the WLAN-first scheme in cellular/WLAN interworking," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1932–1952, May 2007.

[5] M. E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 835–846, Dec. 1997.

[6] S. Thajchayapong and J. M. Peha, "Mobility patterns in microcellular wireless networks," *IEEE Trans. Mobile Comput.*, vol. 5, no. 1, pp. 52–63, Jan. 2006.

[7] H. Zhai, X. Chen, and Y. Fang, "How well can the IEEE 802.11 wireless LAN support quality of service?" *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 3084–3094, Nov. 2005.

[8] H. Zhai, X. Chen, and Y. Fang, "A call admission and rate control scheme for multimedia support over IEEE 802.11 wireless LANs," *Wireless Networks*, vol. 12, no. 4, pp. 451–463, July 2006.

[9] A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long-tail distributions to analyze network performance models," *Performance Evaluation*, vol. 31, no. 3-4, pp. 245–279, Jan. 1998.

[10] N. Benameur, S. B. Fredj, F. Delcoigne, S. Oueslati-Boulahia, and J. W. Roberts, "Integrated admission control for streaming and elastic traffic," in *Proc. 2nd Int'l. Workshop on Quality of Future Internet Services*, Sept. 2001, pp. 69–81.

[11] M. Naghshineh and A. Acampora, "QoS provisioning in micro-cellular networks supporting multiple classes of traffic," *Wireless Networks*, vol. 2, no. 3, pp. 195–203, Aug. 1996.

[12] L. Kleinrock, *Queueing Systems, Volume 1: Theory*. New York: John Wiley and Sons, 1975.

[13] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless Commun. Mobile Computing*, vol. 4, no. 8, pp. 917–931, Dec. 2004.

[14] D. Mitra and J. A. Morrison, "Asymptotic expansions of moments of the waiting time in closed and open processor-sharing systems with multiple job classes," *Adv. Appl. Prob.*, vol. 15, no. 4, pp. 813–839, Dec. 1983.

[15] L. X. Cai, X. Shen, J. W. Mark, L. Cai, and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, no. 3, pp. 752–761, May 2006.

[16] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *Proc. IEEE INFOCOM'04*, vol. 2, Mar. 2004, pp. 1404–1413.

[17] J. Pérez-Romero, O. Sallent, R. Agusti, and M. A. Diaz-Guerra, *Radio Resource Management Strategies in UMTS*. New York: Wiley, 2005.

[18] F. P. Kelly, *Reversibility and Stochastic Networks*. New York: Wiley, 1979.

[19] M. Marsan, "Performance analysis of hierarchical cellular networks with generally distributed call holding times and dwell times," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 248–257, Jan. 2004.

[20] H. L. Vu and T. Sakurai, "Accurate delay distribution for IEEE 802.11," *IEEE Commun. Lett.*, vol. 10, no. 4, pp. 317–319, Apr. 2006.

[21] ANSI/IEEE Std 802.11, "Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 1999.

[22] Maplesoft, "Maple 9.5," http://www.maplesoft.com.

Wei Song (S'07) received the B.S. degree in electrical engineering from Hebei University, China, in 1998 and the M.S. degree in computer science from Beijing University of Posts and Telecommunications, China, in 2001. She is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Her current research interests include resource allocation and quality-of-service (QoS) provisioning for the integrated cellular networks and wireless local area networks (WLANs). She received the Best Student Paper Award from the IEEE Wireless Communications and Networking Conference (WCNC) in 2007.

Yu Cheng (S'01-'04) received the B.E. and M.E. degrees in Electrical Engineering from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2003. From September 2004 to July 2006, he was a postdoctoral research fellow in the Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada. Since August 2006, he has been with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, Illinois, USA, as an Assistant Professor. His research interests include service and application oriented networking, autonomic network management, Internet performance analysis, resource allocation, wireless networks, and wireless/wireline interworking. He received a Postdoctoral Fellowship Award from the Natural Sciences and Engineering Research Council of Canada (NSERC) in 2004 and the Best Paper Award from the International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine) in 2007.

Weihua Zhuang (M'93-SM'01) received the Ph.D. degree in electrical engineering from the University of New Brunswick, Canada. Since October 1993, she has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, where she is a Professor. Dr. Zhuang is a co-author of the textbook *Wireless Communications and Networking* (Prentice Hall, 2003). Her current research interests include multimedia wireless communications, wireless networks, and radio positioning. She received the Premiers Research Excellence Award in 2001 from the Ontario Government for demonstrated excellence of scientific and academic contributions, the Outstanding Performance Award in 2005 and 2006 from the University of Waterloo, the Best Student Paper Award from the IEEE Wireless Communications and Networking Conference (WCNC) in 2007, the Best Paper Award from the IEEE International Conference on Communications (ICC) in 2007, and the Best Paper Award from the International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine) in 2007. She is the Editor-in-Chief of *IEEE Transactions on Vehicular Technology* and an Editor of *IEEE Transactions on Wireless Communications*, *EURASIP Journal on Wireless Communications and Networking*, and *International Journal of Sensor Networks*.